

PEPCY PRACTICAL GUIDELINES (extended)

TITLE: Optimising extraction and HPLC analysis of peptides cyanobacterial strains

AUTHORS: M.Welker

PPG ID: PPG_TUB1

DATE: Aug. 06

PAGES: 10

REVISION LEVEL: 1

1. PURPOSE

The goal of the tailoring procedures is to allow the quantification of peptides in cyanobacterial strains, especially in culture experiments. The following guideline is neither applicable directly to field samples since the peptide composition would be different, nor to preparative procedures since in that case quantification is less important. Results obtained for a particular strain, however, will be a good starting point for others. Without a gravimetric standard only a relative changes in peptide content of a strain can be measured.

2. INTRODUCTION

Cyanobacterial peptides are structurally very diverse and this diversity is also reflected in their behavior when extracted and analysed by high-performance liquid chromatography (HPLC). When an individual strain is studied with respect to the peptide production, an analytical method likely has to be optimised. This requires evaluation of whether any peptide of interest is extracted efficiently and the HPLC-programm has to be optimized to avoid biases due to sub-optimal separation.

With the following guideline it should be possible to develop a HPLC-method specifically optimised for the strain and the peptides under study. If the work is done in a co-operative project the exchange of information and experiences, including the failures, is crucial for the success and can save a lot of time to all partners. At the end of the procedure a HPLC method should stand that has been tested with appropriate statistics for its efficiency and repeatability to meet the standards for publication in peer-reviewed journals.

3. REQUIREMENTS

- HPLC system equipped with a solvent delivery system allowing gradient elution and a photo-diode array detector (PDA)
- analytical reversed phase C18 column (recommended), other columns successfully used for peptide analysis are also applicable
- autosampler; not strictly required, but very helpful to reach the necessary number of replicate runs for appropriate statistics
- SpedVac or similar device for the fast and simultaneous drying of samples and extracts
- Ultrasonic bath for the extraction of cell material
- centrifuge
- cell material of the strain(s) of interest
- reaction tubes
- membrane or glass fiber filters (when respective sampling is desired)

3.1. ACCOMPANYING STUDIES: PEPTIDE IDENTIFICATION

With the present guidelines an identification of peptides is not directly possible. For any strain under study, the peptide profile should be determined individually. It is not safe to rely on retention times and UV-spectra of known peptides to assume these peptides in a strain that has not yet been tested. The best way to do so is the collection of peak fractions and having them analysed off-line, e.g., by MALDI-TOF MS (Czarnecki et al. 2006). In-line methods like LC-MS/MS are helpful, but since the chromatographic system is generally not the same, the results have to be considered carefully.

4. OPTIMISATION OF THE ELUENT PROGRAMM

With this procedure the eluent conditions are optimised to allow the separation of the compounds (peaks) of interest from each other and from background/matrix peaks.

In an ideal case, the compounds of interest elute well separated and can be integrated from baseline to baseline like in Fig. 1a with a peak height between 0.05 and 0.5 AU. Unfortunately this will likely not be the case for many samples. When compounds elute close to each other but the peaks are still baseline separated, the quality of the chromatogram becomes strongly dependent (Fig. 1b) on the amount of the compounds that is loaded on the column (and of column age, of course). In such a case, the amount injected is a trade-off between the detectability and a beginning overlapping of peaks, i.e. between peak height and peak width. Compounds that elute still closer to each other could still be quantified but only with a larger error and only if the peaks have about the same area (Fig. 1c). Critical is the quantification of shoulders and this should thus be avoided (Fig. 1d).

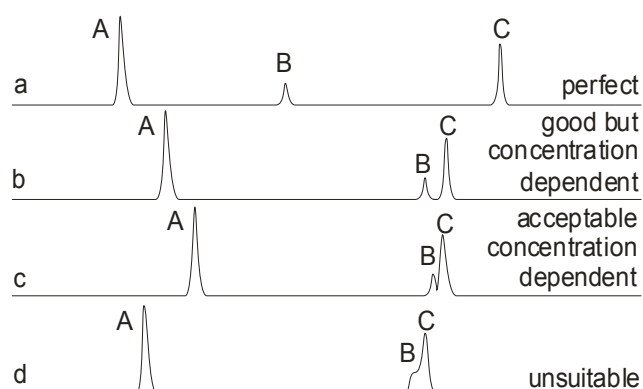


Fig. 1 : Schematic chromatograms of different quality with respect to the quantification of compounds A, B, and C.

4.1 PREPARATIVE STEPS

4.1.1. SELECTION OF AN APPROPRIATE WAVELENGTH

Since most peptides do not have a characteristic UV-absorption spectrum (except microcystins) a wavelength should be chosen where all compounds to be detected absorb well (and preferably other substances much less). Most peptides have a shoulder in the UV-absorption spectra around 220 to 225 nm and therefore a wavelength in that range is suitable for the extraction of chromatograms from 3-dimensional PDA data. If the strain produces peptides with different absorption spectra it might be worth to extract chromatograms at different wavelengths, i.e. 238 for microcystins and 220 for other peptides. It has to be kept in mind that wavelength toward 200 or even 190 nm have a highly increased background absorption for which the software might correct for, resulting in an apparently flat baseline

despite a much higher absolute absorption (see data on the eluent bottles), thus introducing a considerable error in the quantification process.

4.1.2. PROGRAMMING OF ELUENT GRADIENTS

There are several eluent systems that have been applied to rp-C18 columns to detect peptides (microcystins) with success. The most widely applied likely are acidified water/acetonitrile gradients with 0.05% v/v trifluoro-acetic acid. Alternatively eluents buffered with 4 mM ammonium acetate have been used. For the following the use of water/acetonitrile + 0.05% TFA is recommended. For other eluents the same procedures apply in principle.

An HPLC gradient program consists of four phases (Fig. 2):

1. The actual gradient used for the **separation and analysis**; either a single linear increase of the percentage of organic solvent or multiple linear increases with different slopes like the one developed by Lawton et al. (1994) for microcystins.

2. Following the gradient a **purge phase** is necessary to wash off all lipophilic compounds that may clog the column after a while or elute as shadow peaks in later runs. Some 4 to 6 minutes of 100% organic solvent should be sufficient, but this depends on the column dimensions and the composition of the injected extract (see below).

3. After purging with the organic solvent the column has to be **conditioned** to the hydrophilic initial conditions. This has to be done with care especially when the initial conditions are > 80% water. If the change from 100% to 10% organic solvent is too fast the hydrophobic phase (the C₁₈ chains attached to the silica matrix) can collapse. When this happens, in the following run the retention efficiency is strongly reduced resulting in shortened retention times and unexpected chromatograms. To be on the safe side the increase in percentage of water should not exceed 20% min⁻¹. This is a rough estimate and if shifts in retention time were encountered the conditioning phase should be expanded, i.e., made 'shallower'. For very hydrophilic initial condition (>90% water) special RP-columns are available for which the manufacturers claim that these work properly even with 100% water as initial eluent condition.

4. Once conditioned, the column has to be **equilibrated** for about 10 min (or at least one column volume). Properly equilibrated, the baseline will not shift downwards in the first minutes.

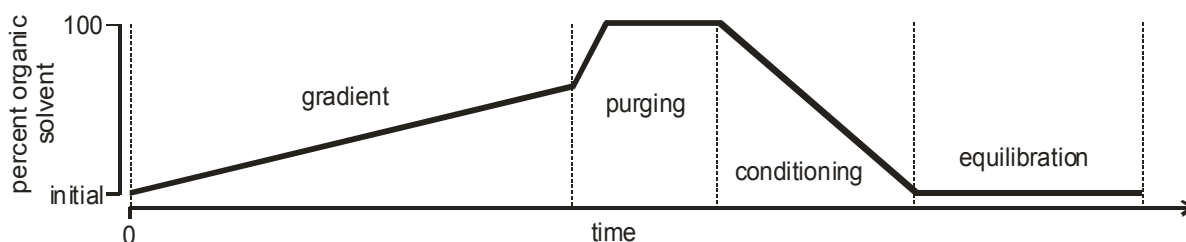


Fig. 2: Schematic illustration of a complete gradient program including the phases to clean and prepare the column for the next injection

4.1.3. PREPARATION OF AN EXTRACT FOR GRADIENT OPTIMISATION

At this point a quantitative extraction is not necessary. Nonetheless, all weights and volumes should be noted to have a first idea about peak areas to be expected from a given

amount of cell material. The easiest is to extract freeze dried biomass three times in 70% methanol. This can be done in Eppendorf tubes. Supernatants are combined and dried completely, e.g., in a SpeedVac.

The dried extract is resolved in 50 % methanol, preferably by first adding one volume of methanol followed by one volume of water. The extract is then allowed to stand for some few hours at 4° to allow all lipophilic compounds to flocculate. Peptides generally are not lost from the liquid phase by doing so. The extract is centrifuged at maximum speed for at least 10 min. and the clear supernatant is taken for HPLC analyses. Care has to be taken to not transfer any particles to the HPLC vial.

The volume of the extract to be injected depends on the biomass extracted and the number of analyses that are intended to be run. A rough calculation should be made before starting. The extract should not be too concentrated (e.g., if 10 µL are injected and absorbance of some peaks reaches 1 AU or more a dilution is recommended)

100 % methanol as for the extract to be injected is strongly disadvised. As a rule of thumb any extract injected to a reversed phase column should be at least as hydrophilic as the initial conditions. All very lipophilic substances like carotenoids (they often give the extracts the colour, not the peptides) will precipitate in 80% water and clog inline filters, precolumns etc. Clarification the extract by centrifugation is recommended rather than filters since little is known about the adsorption of peptides to various membranes.

Dried aliquots of the extract should be stored at -20°C. Likely, peptides are generally stable under these conditions for months or even years. The number of aliquots should be sufficient to have enough for later tests of analytical performance.

4.2.2. RUNNING HPLC ANALYSES

4.2.2.1. SELECTION OF PRACTICABLE ON-COLUMN AMOUNTS

As a starting point, the pumps are programmed for a gradient of 20 to 80 % ACN in 40 min.

After several runs with increasing injection volumes (e.g. 10, 50, 100, 200 µL) it has to be judged whether the peaks of interest are well detected for the following procedures. Further, a first test of linearity in response for well separated peaks can be made. Peak height ideally should be in a range of 0.05 to 0.5 AU. Peak heights of > 1 AU should be avoided since linearity might no longer be given, especially at a wavelength close to 200 nm. Peaks with heights << 0.05 AU can be well detected and quantified but this strongly depends on the detector and the separation.

At this point an estimate of the relative peak height in relation to biomass can be made to get an idea what is needed later when samples will be taken during experiments.

4.2.2.2. EVALUATION OF PEAK PURITY

A good and presumably pure peak should be symmetric (Fig. 3a). This can be measured by measuring a and a' at $h/2$. In the best case $a = a'$, a good indication of peak purity and good general performance. However, often a peak tailing can be observed (Fig. 3b), most likely due to the age of the column, but some substances simply tend to 'tail'. In such a case $a < a'$, but the peak has no shoulders. Shoulders of peaks due to impurity are sometimes not very pronounced and those peaks have to be observed closely in the chromatograms. Some HPLC software packages (i.e. Millennium from waters) have a function of purity check and software manuals should be consulted for this. What is generally done in these procedures is a comparison of the absorption spectra at multiple retention times within the peak (see Fig. 3c). This can

also be done manually and in a peak of a single compound spectra are identical throughout the peak. Another way is to zoom in and open the 3D view of respective peaks, if available. Record spectral data with the highest resolution the system allows, at least < 2 nm in order to get as much information as possible. When $a > a'$ most likely two compounds are eluting close to each other, the one with lower peak area preceding the bigger one.

If the software supports the plotting of the first and second derivative of a peak this might help to detect impurities easily.

But even if all absorption spectra in a (broad) peak are identical this does not mean that it is a single compound! Especially congeners of a single peptide class likely have the same absorption spectra.

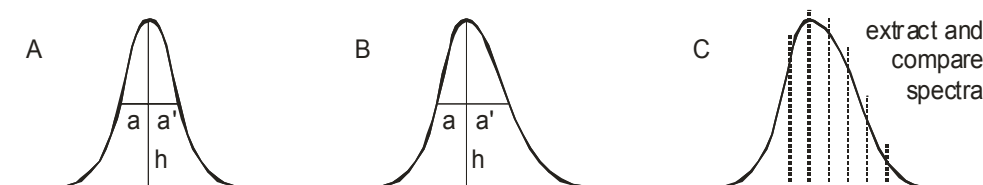


Fig. 3: Ideal (A) and less ideal peaks: a peak with tailing (B) and one with a coeluting compound. h : peak height; a , a' width at $h/2$; dotted lines indicate where spectra should be extracted and compared

In most cases, however, peaks will not be baseline-separated and several compounds will elute closely to the compounds of interest. In that case the integration by the respective software has to be checked and a manual integration has to be performed if necessary. Especially when samples of a wider concentration range were analysed the software settings often result in visibly unprecise intergration when background noise changes. There are no general rules how to do this, but since chromatograms should be very similar from one run to the next, the integration should be similar too, i.e., a baseline starting and ending always with the same peaks.

4.2.2.3. TESTING SEPARATION

Repeated injection (3-5) of the same volume of extract (based on the previous results on suitable on-column amounts) and comparison of the resulting chromatograms:

- a) separation of peaks of interest
- b) peak purity according to 5.2.2.2
- c) classification of results

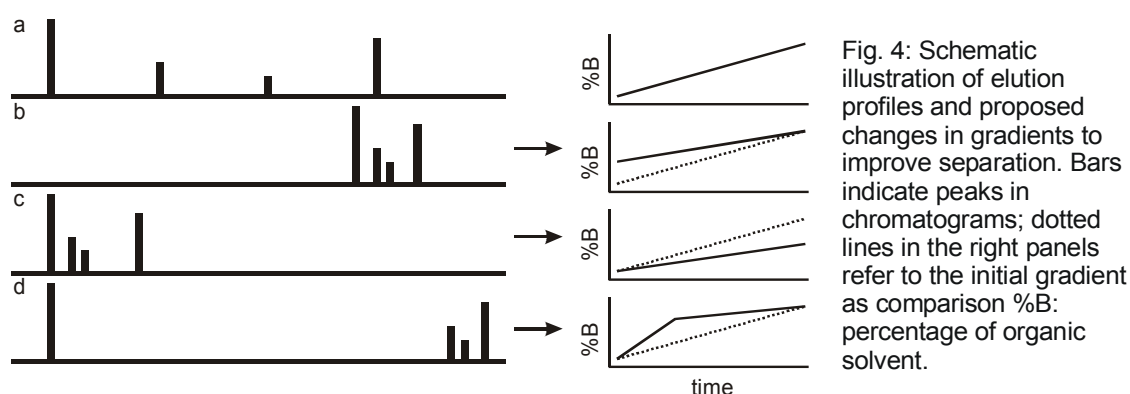
- All peaks are well separated and elute over the full time of the gradient (fig. 4a): perfect
- All peaks elute late (fig. 4b): increase of initial percentage of organic solvent
- All peaks elute early (fig. 4c): decrease of initial percentage of organic solvent
- Some peaks elute very early, others very late with a gap inbetween (fig. 4d): apply a combination of linear gradients.

Whenever the initial conditions are changed, the conditions of the equilibration phase have to be changed accordingly.

The chromatogram has not to be perfect. The gradient should be suitable to separate the peptides that will be quantified, e.g., in culture experiments, and is not intended to be applied to other strains or field samples without modification or evaluation.

If two peaks elute very closely or even co-elute, there is a good probability, that baseline separation will not be achieved by any gradient with the given eluent system. In case a quantification is possible, the best concentration range, i.e. amount on column (o.c.), should be found (see 1). If no satisfying separation can be achieved after several unsuccessful trials a solution could be a change of eluents, e.g. from water/ACN + 0.05% TFA to water/ACN 4 mM ammonium acetate. This will change the hydrophobicity of the compounds differentially due to differences in K_A/K_B values and thus change the relative retention times. This step, however, should be taken only as a last step since it requires intense purging of the entire system.

If an autosampler is available the procedure of gradient optimisation is not too laborious since a sample set can be programmed to be analysed with different gradients. A whole set of samples could then be run automatically, i.e., with gradients 20 to 80% ACN, 30 to 80% ACN, 20 to 70 % ACN etc. A triplicate injection for each gradient is recommended to have numbers for a statistical evaluation at hand and to be sure the results are no artefacts, e.g., due to insufficient equilibration time.



4.2.2.4. TESTING REPEATABILITY

Once suitable eluent conditions have been found and as well as suitable on-column amounts, repeatability has to be checked. To do so, the same volume is repeatedly injected (5-10 times) applying the same gradient (Welker et al. 2002):

- a) check integration, manual re-integration might be necessary
- b) calculation of means, standard deviation, and coefficient of variation ($CV = \frac{SD}{\text{mean}} * 100\%$) for retention time, peak area, and peak height of individual compounds

The CV should not exceed 5 % for all variables. With a well performing system $CV < 1\%$ is possible for most compounds. Special attention has to be paid to outlying values and systematic shifts from the first to the last injection.

For the former error several reasons are possible:

- Conditioning and equilibration time not long enough. In case of doubt both phases in the gradient program should be extended.
- The autosampler does not work properly or the settings are inappropriate. In combination with certain vials and viscous extracts the rate with which the sample is drawn from the

vial can be a source of error. Reduce the speed, especially when using narrow low volume inserts.

- Systematic, directed increases of peak area can occur when evaporation (e.g., with 50 % MeOH as solvent) occurs during the time the samples are placed in the autosampler. The best is to use self-sealing septa if available for the vials. If evaporation is ruled out the age of the column and the gradient have to be checked (see above).

5. OPTIMISATION OF THE EXTRACTION PROCEDURE

Before starting this procedure it has to be made sure that enough standardized material is available, i.e., some 250 mg of DW or some 30 filters loaded with cells from a single harvest. The choice between these two options is made according to the planned procedures for the experiments. As filters either membrane filters (up to 50 mm in diameter, regenerated cellulose) or glass fiber filters have been established for microcystin analysis and there is no reason to doubt their suitability for other peptides. An easier procedure, however, is the simple centrifugation of a culture aliquot, given the cells are not buoyant. The pellet can then be freeze-dried directly or dried in a SpeedVac, where the latter method is preferable to avoid losses.

5.1. PREPARATION OF STANDARDIZED SAMPLES

5.1.1. USING DRY BIOMASS

Cell material is placed in a number of reaction tubes. Some 5 mg most likely are appropriate for most strains and peptides, to be sure refer to the results of 4.1.3. The mass has not to be exactly 5 mg but must be weighted as precisely as possible and the mass carefully noted. It is strongly disadvised to 'correct' the mass in a tube by removing 'excess' biomass because freeze-dried material is easily lost due to electrostatic attraction to the spatula or outer wall of the tubes. If much too much material has been transferred in a tube, start with a new tube.

To minimize errors due to adsorption of peptides in solution always the same brand and type of reaction tubes should be used. Glass tubes surely would be better but are much less comfortable in handling. Note carefully the biomass in each of at least 15-20 tubes and store them cold and dry until use. The number depends actually on the number of variables you wish to test (see below).

5.1.2. USING FILTERS

A culture of the strain under study is used to load filters with aliquots of the same volume. When the volumes of several culture flasks are combined it has to be assured that the content is well homogenized. This is also crucial before every single aliquot is taken; especially with buoyant or sedimenting strains a considerable bias can be introduced even when the culture is allowed to stand onky for some few minutes. Since exactly the same culture will not be available again, the number of samples (filters) that are needed has to be thoroughly calculated, including some spare ones. Wet filters are folded and put into 2 mL reaction tubes, frozen, and freeze-dried if no other sample preparation is envisaged for the experiments. Dry filters are stored cool and dark until used.

5.2. EXTRACTION OF SAMPLES

As a starting point an extraction with 70 % methanol is advised since most peptides are extracted to some degree with this solvent. When using 2 mL reaction tubes 1-1.5 mL 70%

MeOH are applied to the cells/filters, the tube content is then mixed and sonicated for 10 min. High temperatures (>30°C) should be avoided since there is the possibility of methyl-ester formation with free carboxy-groups. The samples can be agitated for some time, put in the fridge for some hours, or centrifuged directly. Whatever the procedure is, notes should be taken carefully and all samples should be treated in the same way at this stage. After centrifugation, the supernatant is collected in another tube and the pellet resuspended again in 70% MeOH, sonicated etc.

The combined supernatants are dried completely to avoid decay. Preferably a SpeedVac is used since it allows the efficient and simultaneous drying of some dozens of samples and minimises the risk of loss. For further procedures see 4.1.3.

What is to be tested is the number of extraction steps that has to be performed to extract a given compound completely from the cell material (Fig. 5). Completely in this sense does not mean that all the peptide has to be removed from the cell debris, it means that an accurate measurement of the peptide content of the cells (cell quota) can be performed. An easily extractable compound, for example, which is efficiently extracted with a single extraction step (i.e., by quickly reaching an equilibrium of intra- and extracellular concentration) the pellet, of course, contains still the peptide and following extraction steps will show this. This allows, however, no direct conclusion on the extraction efficiency.

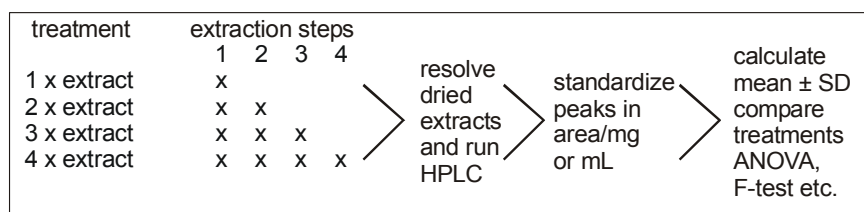


Fig. 5: Scheme for estimation of the number of required extraction steps.

The most accurate way is to perform the extraction with multiple steps on individual samples, i.e., one series of samples extracted once, one series extracted twice, and so on.

As a shortcut, a single sample can be extracted repeatedly and the cell quota be calculated that result from each extraction step. This procedure involves some calculations that are best done in a program like Excel. As a crucial step, the volume of supernatant you draw from the tubes after each extraction step has to be noted carefully. The residual (the solvent bound to cell debris) contains the same concentration as the supernatant and this has to be considered in the calculations. For the next extraction step the concentration has to be calculated which results from the dilution of the residual volume with fresh solvent and compared to the concentration (i.e., peak area) actually measured in the second supernatant and so on. The variable that is used to judge the efficiency is the peptide content of the cell. Any concentrations measured as intermediate variables have to be converted to it. A particular extraction step is efficient when the calculated cell quota increase by this extraction step.

5.3. STATISTICAL EVALUATION OF EXTRACTION EFFICIENCY

A reasonable range of extraction steps to test is from 1 to 4 since an extraction method should not require more than 3 extraction steps for practical reasons. Respective extracts are prepared and injected to HPLC. The peak areas of individual compounds have to be standardized to give a value of peak area mg^{-1} DW (i.e., cell quota) or peak area mL^{-1} culture volume, respectively.

The best way to do some basic statistics is by putting the standardized peak area data in an Excel spread sheet followed by calculations of means, SD's, and CV's for each peptide and

extraction procedure separately from the three replicates. The extraction of an individual compound can be assumed as efficient, when there is no statistically significant increase in calculated cell quota from one step to the next. This can be tested simply by comparing the data in a plot, for example. A precise way would be to perform statistical calculations such as Student's test or F-test or to run an ANOVA followed by an appropriate test. It has to be ensured why a certain extraction procedure is considered as efficient and a statistical evaluation is very helpful in this context.

A plot might look like Fig. 6. Compound A is extracted very efficiently and after the second extraction no further increase in relative peak area is observed, i.e. two extraction steps are sufficient. The extraction efficiency for compound C is very low, the fourth extraction step results in a significant increase of calculated cell quota and thus the number of required extraction steps is at least five. For compound B a number of three extraction steps is sufficient to extract 100 % from the cells.

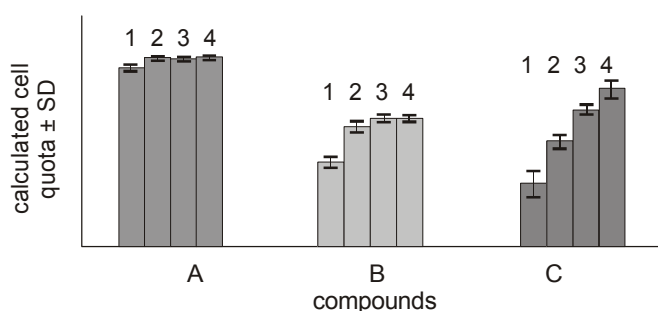


Fig. 6: Exemplary plot of peak area vs. number of extraction steps of three compounds that are extracted with differing efficiency.

An extraction method has then to be judged with respect to the worst extractable compound of interest. Should the number of extraction steps exceed three, it is strongly recommended to make a change in the extraction solvent composition.

To which side the change should go – more hydrophilic or more lipophilic – depends on which compounds are actually extracted unsatisfactory with three extraction steps: if the compound is eluting late, the extraction should be more lipophilic and if eluting early the extraction should be done with a solvent that is more hydrophilic. In the first case, an extraction with 90% MeOH could be tested, in the latter one with 50% MeOH. It should nonetheless be considered that the more lipophilic the extraction solvent is, the more other lipophilic compounds will be extracted (carotenoids, fatty acids, etc.) that may interfere with peptide detection in chromatograms.

Another option is to perform a sequential extraction with, for example, first water, followed by 50% MeOH, and finally 100% methanol.

6. CALIBRATION AND GENERAL PERFORMANCE TESTING

To assess general performance it is recommended to run an analysis of the standard extract (4.1.3) from time to time to check retention times, spectra, and baseline drift. A relative calibration can be performed by calibrating the system with an external standard, e.g. microcystin-LR, and to check the ratio of peak area to amount on column regularly. The ratio should fit well in the respective calibration curve that has been established previously. In that case, the lamp and the detector are performing well and all other peak areas can be used to calculate peak areas relative to the extracted biomass.

If the peak area to amount on column ratio of the external standards deviates from the calibration curve, the performance of the detector has to be checked. Most systems have in-

built diagnostics programs to do so. Further the age of the lamp has to be checked: toward the end of the lamp life-time the emitted light is getting weaker resulting in less clear spectra and chromatograms. Before discarding the –generally very expensive – lamp a thorough flushing of the flow cell is recommended with 100% MeOH, 100% water, and again 100% MeOH after the column has been removed.

The use of an internal standard is also an option. This means to add microcystin-LR in a known amount to the extracts and compare the peak area (in relation to the injected volume) from one run to the next. If the system performs properly the variation (CV) should be in a range as determined previously (4.2.2.4). For many strains this will, however, not be applicable since the strains themselves produce microcystins. Further, it might be not desired to add a new peak to chromatograms that might interfere with peaks of interest.

A quantitative analysis, however, requires true gravimetric standards. As soon as pure compounds are available as standards a direct calibration can be undertaken. This can also be done, though not absolutely correct, on samples analyses before, given the system was regularly checked for good performance and repeatability.

7. REFERENCES

- Czarnecki O, Lippert I, Henning M, Welker M (2006) Identification of peptide metabolites of *Microcystis* (Cyanobacteria) that inhibit trypsin-like activity in planktonic herbivorous *Daphnia* (Cladocera). *Environ.Microbiol.* 8:77-87
- Lawton LA, Edwards C, Codd GA (1994) Extraction and High-Performance Liquid Chromatography method for the determination of microcystins in raw and treated waters. *Analyst* 119:1525-1530
- Welker M, Bickel H, Fastner J (2002) HPLC-PDA detection of cylindrospermopsin - opportunities and limits. *Wat.Res.* 36:4659-4663